THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

Theoretical Characterization of the H-Bonding and Stacking Potential of Two Non-Standard Nucleobases Expanding the Genetic Alphabet

Mohit Chawla, Raffaele Credendino, Edrisse Chermak, Romina Oliva, and Luigi Cavallo

J. Phys. Chem. B, **Just Accepted Manuscript** • DOI: 10.1021/acs.jpcb.6b00125 • Publication Date (Web): 16 Feb 2016

Just Accepted

ACS Publications

The Journal of Physical Chemistry B is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Theoretical Characterization of the H-Bonding and Stacking Potential of Two Non-Standard Nucleobases Expanding the Genetic Alphabet

Mohit Chawla,[1,*] Raffaele Credendino,[1] Edrisse Chermak[1] Romina Oliva,[2] Luigi Cavallo[1,*]

[1]King Abdullah University of Science and Technology (KAUST), Physical Sciences and Engineering Division (PSE), Kaust Catalysis Center (KCC), Thuwal 23955-6900, Saudi Arabia. [2]Department of Sciences and Technologies, University Parthenope of Naples, Centro Direzionale Isola C4, I-80143, Naples, Italy. Email: mohitchawla.bt@gmail.com; luigi.cavallo@kaust.edu.sa

Abstract:

We report a quantum chemical characterization of the non-natural (synthetic) H-bonded base pair formed by 6-amino-5-nitro-2(1H)-pyridone (Z) and 2-amino-imidazo [1,2-a]-1,3,5-triazin-4(8H)-one (P). The Z:P base pair, orthogonal to the classical G:C base pair, has been introduced in DNA molecules for expanding the genetic code. Our results indicate that the Z:P base pair closely mimics the G:C base pair both in terms of structure and stability. To clarify the role of the NO_2 group on the C5 position of the Z base, we compared the stability of the Z:P base pair with that of base pairs having different functional groups at the C5 position of Z. Our results indicate that the electron donating/withdrawing properties of the group on C5 has a clear impact on the stability of the Z:P base pair, with the strong electron withdrawing nitro group achieving the largest stabilizing effect on the H-bonding interaction, and the strong electron donating NH_2 group destabilizing the Z:P pair by almost 4 kcal/mol. Finally, our gas phase and in water calculations confirm that the Z-nitro group reinforces the stacking interaction with its adjacent purine or pyrimidine ring.

Introduction:

Synthetic Biology is an emerging research area, and its key objective is to design novel biological systems by introducing newly designed non-natural components.[1-3] Within this approach, the newly designed non-natural components work alongside with natural components in a new biological system.[1,3] In this scenario, during the past few decades synthetic biologists have put remarkable efforts to create non-natural base pairs that function toward the expansion of the genetic alphabet of DNA. The genetic information encoded in DNA is stored using a four letters alphabet, namely Adenine (A), Thymine (T), Guanine (G) and Cytosine (C), and in a double helical DNA structure 'A' selectively pairs with 'T' and 'G' pairs with 'C'. This pairing is fundamental to the transmission of the genetic information through replication, transcription and translation. Thus, introducing a non-natural alphabet allows increasing the genetic alphabet, resulting in the expansion of the genetic information that can be encoded by DNA. Apart from the notion of increase in genetic information, it is worth to point out that non-natural base pairs have been studied extensively for various applications, including the site specific labeling,[4-6] specific detection/probing,[7,8] immobilization[9] and for the structural analysis of nucleic acids.[10-12]

During the last 30 years, a series of nucleotide analogues for the expansion of the genetic code have been reported by the groups of Kool, Hirao, Benner and Romesberg,[5,6,9-27] mainly. An ideal feature of the non-natural nucleotides introduced in DNA is their recognition by the natural enzymes machinery, which should be able to replicate it, as well as to transcribe the non-natural DNA into a non-natural RNA containing extra ribonucleotides. This would allow increasing the number of codons that, in turn, could encode proteins containing non-natural amino acids. Most of the reported non-natural nucleotides[11,28-30] lack the capability to form H-bonded pairs with other nucleotides, and deviate substantially from the concept of Watson-Crick base pairing. For example, non-standard nucleobases d5SICS and dNaM intercalate in duplex DNA, rather than lying coplanar, which leads to significant distortion of a regular double helical structure of DNA, and the edge-on interaction is only enforced when duplex DNA interacts with polymerase.[28-30]

This weakness of the single non-natural nucleotide can be overcome through the development of biorthogonal non-natural base pairs. For example the Z:P base pair, composed by 6-amino-5-nitro-2(1H)-pyridone (Z) and 2-amino-imidazo [1,2-a]-1,3,5-triazin-4(8H)-one (P), see Figure 1, which combines with the four naturally occurring DNA bases to form a six letters DNA alphabet.[31,32] The Z:P pairing is orthogonal to pairing in standard bases, as only the hydrogen bonding units are

shuffled.[31,32] It is to be remarked that DNA containing these Z and P non-natural bases can be amplified by polymerase with very high fidelity (99.8% per cycle).[31,32] This fidelity is due to the fact that Z and P nucleotides form natural stacking interaction, which is highly similar to the ones adopted by Watson-Crick pairs.[31,32] This allows the incorporation of multiple non-natural base pairs adjacent to one another, without perturbing the fundamental DNA duplex structure.[32] Moreover, it has also been shown that Z:P-containing DNA can adopt both the A- and the classical B-form structures.[32] Finally, a RNA riboswitch with a Z:P pair replacing a C:G pair has also been recently shown to preserve its conformation and ligand affinity.[33] From the structural point of view, the minor groove of the Z:P base pair is very similar to that of natural A:T and G:C base pairs, thus these non-natural base pairs could efficiently interact with polymerases.[31,32] A unique feature associated with Z:P containing DNA is imparted by the Z-nitro group present at the major groove of the Z base, which could potentially be exploited for recognition by proteins.[31,32] Moreover, it was suggested that the strong electron withdrawing nitro group stabilizes the nucleoside derivatives against epimerization,[34] and that it could influence both the H-bonding and the stacking potential in the DNA structure.

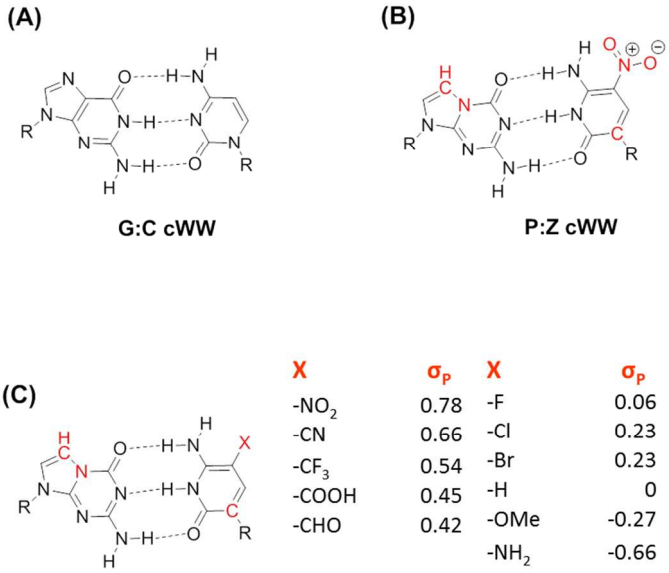


Figure 1. Structure of (a) G-C base pair; (b) Z-P base pair; (c) Z-P base pair with different functionalities (X) introduced at C5 position of Z base, with the values of the Hammett constant characterizing each substituent.

Nevertheless, all these considerations lack a formal quantification based on electronic structure methods, which proved extremely useful to characterize bases pairing and stacking in natural nucleic acids.[35-50] To fill this gap and to complement the reported structural studies,[32] we decided to perform quantum mechanics calculations. Following an approach we used to evaluate the impact of natural and non-natural modifications on H-bonding base pairing,[36,38] we initially focused on the effect of the modification on the geometry and energetics of the base pair, and compared them with the standard G:C Watson Crick pair. The final goal is to understand to which extent the non-natural Z:P pair differs from the G:C pair, since the H-bond pairing has to respect stability criteria, balancing between a too weak pairing, that could induce instability in the non-natural DNA, and a too strong pairing, that could result in difficult processability by the enzyme machinery.

Furthermore, to have a better understanding of the dependence of the H-bonding capability of the Z:P base pair on the nature of the substituent at the C5 position, we also analyzed the modified Z base bearing functional groups possessing different electronic properties, which can be quantified by the values of the Hammett constant, σ_P.[51] The first two functional groups we studied are the cyano (-CN) and the trimethyl-fluoride (-CF_3) groups, which are strong electron withdrawing as the nitro group, and have been previously studied extensively for different applications in context of nucleic acid structures.[52-55] In particular, the 5-CN-Z:P base pair has been designed to study the influence of a different electron withdrawing group (cyano versus nitro) on epimerization and enzymatic incorporation of substituted Z:P pair.[56] Next, we focused on the formyl (-CHO) and carboxyl (-COOH) groups, which are much less electron withdrawing and that also preserve the intramolecular H-bond, similar to the one observed between the amino and the nitro group in the parent Z base, see Figure 2. These formyl and carboxyl moieties have also been very well studied in case of epigenetic modifications, in particular C5 atom modification of cytosine in the context of DNA structures.[57,58] Next, we focused on the electron donating methoxy (-OMe) and amino (-NH_2) group at the C5 position, as -NO_2 and NH_2 are at the extremes of the Hammett scale,[51] which measures the electron donor and withdrawing capability of substituents. This will allow us to understand to which extent the electronics and the H-bonding properties of the Z base can be tuned. Finally, the analysis is completed by considering the impact of a halide substituent, F, Cl and Br, on the C5 atom of the Z base, again to examine the influence of functional groups that can differently impact the electron density on the base. For all these systems we considered the classic Watson-Crick (cWW) geometry, and studied the systems both in gas phase and in water. As a final remark, we also studied the modeled Z:P tWW base pair, that can be accommodated in parallel stranded nucleic acid structures.

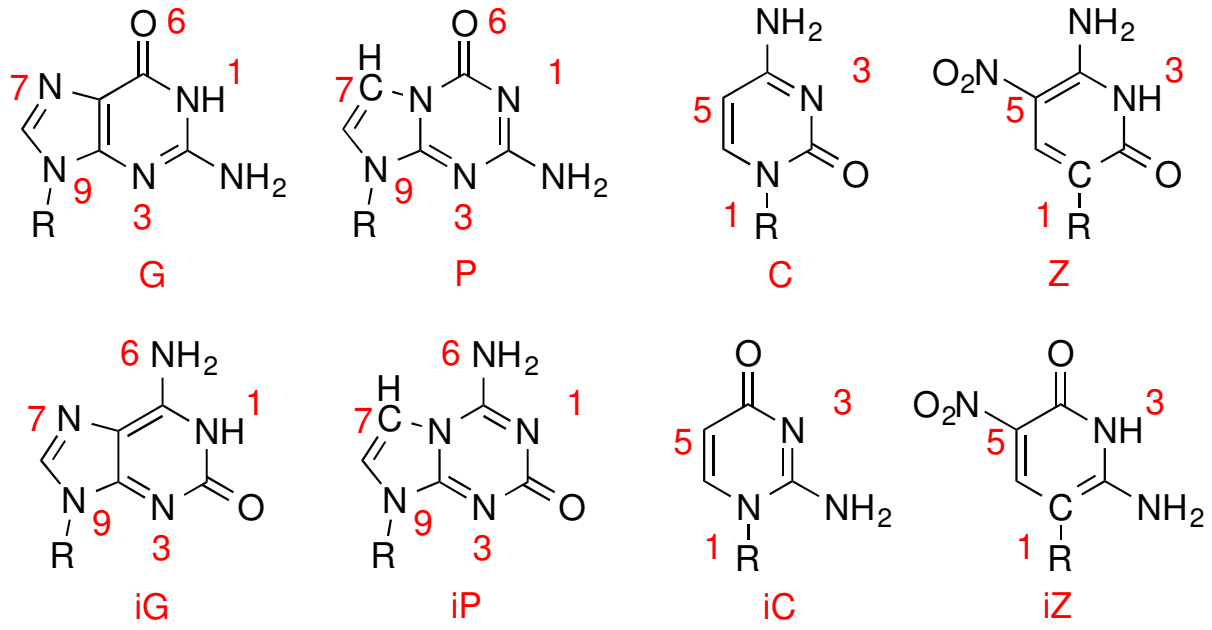


Figure 2. Structure of: Guanine, G; 2-amino-imidazo [1,2-a]-1,3,5-triazin-4(8H)-one, P; Isoguanine, iG; Iso-P, iP; Cytosine, C; 6-amino-5-nitro-2(1H)-pyridone, Z; Isocytosine, iC; Iso-Z, iZ.

Models and Computational Details.

To investigate the stability of the non-natural modified base pairs under study, we have modeled the H-bonded modified base pairs and calculated their interaction energy using density functional theory and post-HF methods.

Modeling the interaction system. We first focused on studying the impact on H-bonding of a novel Z:P pair compared to the classical G:C pair. For this, we just took into account the 6-amino-5-nitro-2(1H)-pyridone heterocycle (Z) H-bonded to the 2-amino-imidazo[1,2-a]-1,3,5-triazin-4(8H)one heterocycle (P) (i.e. Z:P cWW). Next, we studied the impact of different C5 substituents on the 'Z' moiety that is H-bonded to 'P'. Next, to study the impact of the electron withdrawing nitro group on the H-bonding stability of Z:P base pair, we replaced the nitro group with a hydrogen, forming 5-H-Z:P base pair. Finally, the effect on the C5 position of the Z-moiety of different substituents, namely -CN, -CF_3, -COOH, -CHO, -OMe, –NH_2, -F, -Cl and -Br, was also studied (see Figure 1) giving rise to the 5-CN-Z:P, 5-CF_3-Z:P, 5-COOH-Z:P, 5-CHO-Z:P, 5-OCH_3-Z:P, 5-NH_2-Z:P, 5-F-Z:P, 5-Cl-Z:P, 5-Br-Z:P base pairs (see Figure 3). The coordinates of the modeled geometries of the base pair

corresponding to G:C cWW and Z:P cWW were built from the crystallographic structure with PDB ID: 4XNO.[59] For other base pairs, respective functionalities were introduced at the C5 position of the 'Z' base. Thus, in total, we modeled 10 different combinations of the non-natural modified base pair systems. The glycosidic bonds of the described base pairs are disposed in the 'cis' orientation, which corresponds to the geometry pertinent to antiparallel stranded (aps) DNA structure.

To study the impact of the Z:P pair on the parallel stranded nucleic acid structures, in addition to the above described calculations, we also modeled the Z:P base pair considering a 'trans' orientation of the glycosidic bonds. It has to be remarked that, for building the parallel stranded nucleic acid, isoguanine-cytosine and/or guanine-isocytosine pairs have been used instead of the standard guanine-cytosine pair.[60-64] Therefore, for similarity to isoguanine (iG) and isocytosine (iC), we modeled isoP and isoZ, where the carbonyl and amino groups have been flipped with respect to the geometry of 'P' and 'Z' respectively, see Figure 2. Thus, two base pairs were considered. The former being iZ:P tWW, the latter being iP:Z tWW. Finally, to study the impact of the nitro group present on the C5 atom of Z, we replaced the nitro group with a hydrogen atom. In total, 6 modified base pairing combinations that can be accommodated in parallel stranded nucleic acids were modeled. For all the model systems described above, the base pairs are truncated at the C1' atom of the ribose. This is the standard approach used in literature.

QM calculations. Geometry optimizations were performed within a density functional theory approach, based on the hybrid B3LYP functional,[65-67] as implemented in the Gaussian09 package. The correlation-consistent polarized valence triple-ζ cc-pVTZ basis set[68] was used for all the geometry optimizations in gas phase as well as in water, modeled with the C-PCM continuum solvation model.[69] Since dispersion interactions might contribute differently to the stability of the base pairs under study, we also added Grimme D3 correction term to the electronic energy.[70] Interaction energies were calculated on the B3LYP-D3/cc-pVTZ optimized geometries at the second order Møller-Plesset (MP2)[71] level of theory using the augmented aug-cc-pVTZ basis set. For these calculations, we took advantage of the faster RIMP2[72] method as implemented in Turbomole 6.1 package, with water modeled with the continuum solvation model COSMO.[69] All the interaction energies were corrected for the basis set super position error (BSSE),[73] using the counterpoise procedure.[73] Thus, the interaction energy E_{Int} is calculated as in Eq. 1:

$$E_{Int} = E_{Complex} - (E_{M1} + E_{M2}) + \text{BSSE}; \quad (1)$$

where $E_{Complex}$ is the electronic energies of the optimized M1:M2 base pair, and E_{M1} and E_{M2} are the electronic energy of the isolated geometries of the M1 and M2 bases, and BSSE is the basis set superposition error. Within this approach, the deformation energy, which is the energy required to deform the bases from the isolated geometry to the geometry they have in the base pair, is included in our calculations. This is a rather standard approach used in this kind of calculations.[35,39,41,44,48-50] In the present study, we also derived the interaction energies in water, which were calculated using the same recipe as suggested by Sponer and coworkers.[48,74]

To have an immediate and intuitive understanding of the impact of a specific modification, we introduced the modification energy, E_{Mod},[36,38] defined as the energy difference between the interaction energy of the modified and of the corresponding natural base pair (in this specific case Z:P base pair), as shown in Eq. 2.

$$E_{Mod} = E_{Int}(\text{modified base pair}) - E_{Int}(\text{natural base pair}). \quad (2)$$

Within this definition, positive and negative E_{Mod} values indicate modifications that decrease or increase the stability of a specific base pair, respectively.

Electron density analysis. Comparative analysis of the electron density of the Z:P base pair and of the C5-substituted Z base in 5-X-Z:P modified base pairs was performed as follows. First, the geometry of the Z-P base pair was optimized at the B3LYP/cc-pVTZ level of theory. For the sake of easier analysis,[42] C_S symmetry, with the symmetry plane coincident with the Z:P base pair plane, was imposed to the systems, and the electron density analysis was performed in the symmetry plane. After optimization, we compared the RI-MP2/aug-cc-pVTZ electron densities of the modified 5X-Z:P base pair, $\rho^{5\text{-x-ZP}}$, and that of the non-modified Z:P base pair with the geometry it has in the 5-X-Z:P base pair , $\rho^{ZP/5\text{-x-ZP}}$. In other words, we took the optimized geometry of the 5-X-Z:P base pair, and we replaced nitro group present at C5 with different functional groups, referred as 'X' Further, the functional group 'X' was optimized, freezing the coordinates of the common atoms as in natural Z:P base pair. With this approach, the heavy atom skeleton of the 5-X-Z:P and Z:P base pair is identical and can be perfectly superimposed. This is fundamental to avoid noise in the analysis of the electron density difference, $\rho^{5\text{-x-ZP}-ZP/5\text{-X-ZP}}=\rho^{5\text{-X-ZP}}-\rho^{ZP/5\text{-x-ZP}}$.

Evaluation of Stacking Interactions:

We evaluated the base-base stacking interaction energy as described in equation (1), using the DLPNO-CCSD(T) method with cc-PVTZ basis set and BSSE correction. We kept default parameters as implemented in the Orca suite of programs.

Results and Discussion:

Ten different modifications of 6-amino-5-nitro-2(1H)-pyridone heterocycle (Z) on the 'Hoogsteen edge' have been investigated, all of them being derivatives of Z. Optimal geometries and energetics of the base pairs they form with 2-amino-imidazo[1,2-a]-1,3,5-triazin-4(8H)one heterocycle (P), in a classical '*cis*' Watson-Crick (cWW) geometry, have been calculated in the gas phase and in water. Table 1 summarizes calculated energies for the investigated base pairs. Optimal geometry for the *cis* Watson-Crick base pairs is shown in Figure 3, where H-bonding distances in gas and in water are also shown.

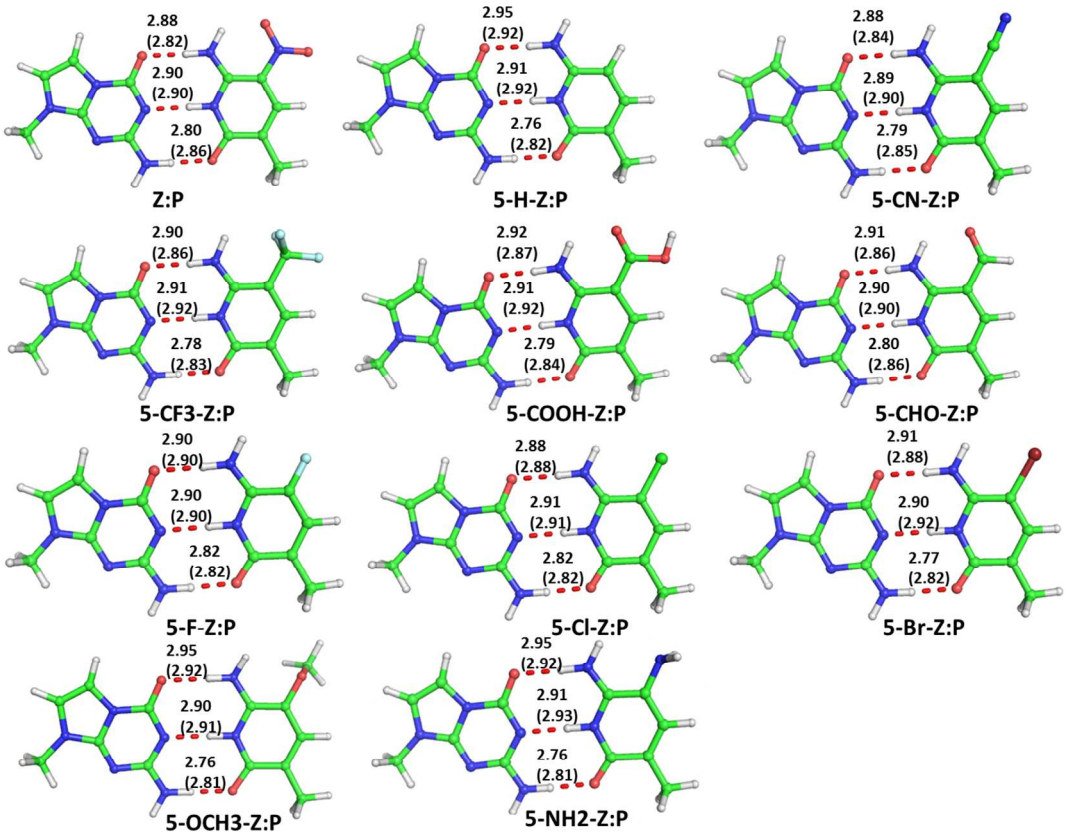


Figure 3. Stick representation of the base pairs including a modified 'Z' base H-bonded to 'P' in the cWW geometry. The values in parenthesis correspond to the optimized distances in water and values without parenthesis correspond to optimized distances in the gas phase. All distances are in Å.

Structural and Energetic Comparison between Z:P cWW and G:C cWW base pair.

The geometry of non-natural Z:P cWW pair is orthogonal to standard G:C pair, as the hydrogen bonding units are only shuffled. Focusing on the overall geometry, the optimized Z:P cWW base pair is almost similar to the classical G:C cWW pair in terms of H-bonding, with differences in H-bond lengths within 0.02 Å. Similarly, the C1'-C1' distance (which gives an indication of the isostericity of H-bonded base pairs),[75,76] of the optimized geometries of Z:P and G:C pairs is almost identical (10.78 Å versus 10.71 Å, respectively). Remarkably, the interaction strength of the Z:P cWW pair is extremely similar to that of the G:C cWW pair, with E_{Mod} of 0.21 kcal/mol and -0.19 kcal/mol in the gas phase and in water, respectively, clearly indicating the comparable stability of the Z:P versus the G:C base pair.

A hypothetical geometry of the Z:P base pair in the tWW orientation has also been studied, which corresponds to a geometry pertinent to parallel stranded nucleic acid molecules. For construction of a parallel stranded nucleic acid, isoguanine-cytosine (iG:C) and/or guanine-isocytosine (G:iC) pairs have been employed instead of the standard guanine-cytosine (G:C) base pair.[60-64] Similarly to isoguanine (iG) and isocytosine (iC), we modeled isoP and isoZ, where the carbonyl and amino groups have been flipped with respect to the geometry of 'P' and 'Z' respectively, see Figure 2. Thus, two possible base pairs were considered: iZ:P tWW and iP:Z tWW.

The geometries of iG:C and iC:G are very similar to Z:iP and iZ:P, respectively, with the differences in optimized H-bonds within 0.15 Å. Moving to the interaction strengths, the modeled Z:iP tWW geometry is slightly more stable than the iG:C tWW geometry with E_{Mod} of – 0.72 and -0.77 kcal/mol in gas phase and in water respectively. In contrast, the P:iZ tWW base pair is marginally less stable than G:iC tWW, with an E_{Mod} of +0.05 and +0.35 kcal/mol in gas phase and in water. The geometric and energetics comparison of the above base pairs clearly points out the comparable stability of the Z:P pairing, thus reinforcing the experimental evidence that the Z:P base pair can potentially fit into a duplex DNA without perturbing its structure, nicely mimicking the G:C base pair in terms of structure and stability. This is an important point, since it indicates how well the Z:P base pair matches the fundamental requirements for being processed by the enzyme machinery.

Table 1. Geometry and Interaction energy values for the cWW and tWW Z:P and 5-substituted Z in Z:P base pairs. All interaction energy values are reported in kcal/mol. E_{mod} is the difference between the interaction energy of the modified base pair and of the reference pair (that is Z:P cWW for the cWW geometries, and iZ:P tWW and Z:iP tWW for the tWW geometries). Negative and positive values of E_{Mod} indicate that the modified base pair is more stable or less stable than the A:U base pair, respectively.

System	Geometry	E_{int}	E_{def}	ΔEtot-gas	E_{mod} (gas)	ΔE-water	E_{-mod}(water)
G:C	cWW	-30.74	2.75	-27.99	-----	-12.49	-----
Z:P	cWW	-30.26	2.48	-27.78	0.00	-12.68	0.00
5-H-Z:P	cWW	-28.92	3.59	-25.33	2.45	-11.69	0.99
5-CN-Z:P	cWW	-30.57	2.58	-27.99	-0.21	-12.54	0.14
5-CF3-Z:P	cWW	-30.21	2.64	-27.57	0.21	-12.43	0.25
5-CHO-Z:P	cWW	-29.03	2.52	-26.52	1.26	-12.11	0.57
5-COOH-Z:P	cWW	-28.81	2.54	-26.26	1.52	-12.02	0.66
5-F-Z:P	cWW	-30.06	4.26	-25.79	1.99	-12.36	0.32
5-Cl-Z:P	cWW	-30.03	4.08	-25.94	1.84	-12.39	0.29
5-Br-Z:P	cWW	-29.85	3.60	-26.25	1.53	-12.34	0.34
5-OCH3-Z:P	cWW	-29.10	3.50	-25.59	2.19	-12.06	0.62
5-NH2-Z:P	cWW	-28.34	3.68	-24.66	3.12	-11.90	0.78
iG:C	tWW	-34.09	3.18	-30.90	----	-13.77	---
Z:iP	tWW	-35.70	4.08	-31.62	0.00	-14.54	0.00
5-H-Z:iP	tWW	-35.77	6.16	-29.61	2.01	-13.80	0.74
G:iC	tWW	-33.96	4.02	-29.94	----	-13.25	----
iZ:P	tWW	-32.84	2.95	-29.89	0.00	-12.90	0.00
5-H-iZ:P	tWW	-30.13	3.16	-26.96	2.93	-12.11	0.80

Studying the Impact of Different Substituents on the C5 atom of 'Z' base in Z:P base pair

A unique feature associated with Z:P containing DNA is imparted by the nitro group present on the C5 atom in the major groove of the Z base.[32] To study the impact of the nitro group on the Z:P pair, we modeled a series of substituents at the C5 site of the Z base in place of the nitro group, see Figure 1. The first geometry corresponds to the Z:P pair with the nitro group on the C5 atom of the Z base replaced by a simple hydrogen atom. We refer to this geometry as 5-H-Z:P base pair. The optimized geometry of 5-H-Z:P is quite similar to that of Z:P, with H-bond differences within 0.10 Å. As for the interaction energy, the absence of the nitro group in the 5-H-Z:P base pair substantially destabilizes the H-bonding strength with E_{mod} values relative to the Z:P pair of +2.45 and +0.99 kcal/mol in gas phase and in water, and of +2.66 and +0.80 kcal/mol relative to G:C. To shed light on the decreased stability of the 5-H-Z:P base pair compared to the Z:P pair, we compared the

electron densities of both the systems. Figure 4 reports the difference of the electron density in the gas phase between the Z:P and the 5-H-Z:P base pairs. Inspection of Figure 4a clearly indicates that the nitro group at the C5 position substantially decreases electron density on groups at the ortho and para positions, i.e. around the O2 and N4 atoms of Z. This phenomenon, which can be easily explained in terms of resonance formulas involving the π molecular orbitals scheme in the 6-membered aromatic ring of Z, has a dual effect. On one side, it reduces the H-bonding accepting ability of the O2 atom, which weakens the bases paring, while on the other side, it makes the N4 amino group a stronger H-bond donor. As for the N3 group in the meta position, inductive effects through the σ-bonds of the 6-membered aromatic ring of Z make it a better H-bond donor. Support to this analysis is also given by a comparison of the H-bond distances in the 5-H-Z:P and Z:P pairs. In fact, the H-bond involving the N4 group of Z is shortened in presence of the nitro group, whereas the H-bond involving the O2 group of Z is elongated in presence of the nitro group. The balance of these effects results in an overall stabilizing effect of the nitro group.

A similar analysis was also performed on tWW orientation of base pairs, which could be employed for construction of the parallel stranded DNA molecules. The optimized geometries of 5-H-Z:iP and 5-H-iZ:P tWW pairs are very similar to those of the Z:iP and iP:Z pairs, with H-bond differences within 0.08 Å. Again, the nitro group has a stabilizing effect, since the E_{mod} values of both 5-H-Z:iP and 5-H-iZ:P are +2.01 and +2.93 kcal/mol in gas phase and +0.74 and +0.80 kcal/mol in water.

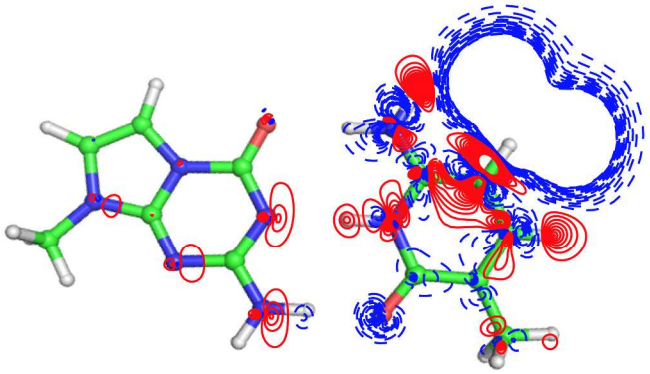


Figure 4. Electron density difference, in the base plane, between the Z:P base pair presenting a nitro substituent on the C5 atom, and the 5-H-Z:P base pair. Density difference curves are plotted between –0.02 and 0.02 a.u., with a spacing of 0.001 a.u. Blue (red) lines refer to negative (positive) density difference curves, i.e., to areas where the base pair including the substituted 'Z' presents reduced (increased) electron density as compared to the Z:P base pair.

Intrigued by the above results, which clearly indicate the non-innocent role of the substituent on the C5 position of Z, we modeled a series of different substituents, with varying electronic properties, at the C5 position of the Z base. The first two functional groups we studied are the cyano (-CN) and tri-methyl-fluoride (-CF_3), resulting in the 5-CN-Z:P and 5-CF_3-Z:P base pairs. The cyano and tri-fluoro-methyl functional groups possess similar electronic properties to that of nitro group (i.e. strongly electron withdrawing), and have been extensively studied previously for different applications in the context of nucleic acid structures.[52-55] As expected, the geometry of the 5-CN-Z:P and 5-CF_3-Z:P base pairs is very similar to that of the Z:P base pair, with H-bond differences within 0.04 Å. As for the interaction energies, the 5-CN-Z:P and 5-CF_3-Z:P base pairs have a comparable stability, with E_{mod} of -0.21 and +0.21 kcal/mol in gas phase and of +0.14 and +0.25 kcal/mol in water, respectively, as compared to the Z:P base pair. This confirms that a strong electron withdrawing group on the C5 atom of the Z base is essential to tune the energy of the Z:P pair on values comparable to the classical G:C pair in the DNA structure, see Table 1.

Next, we focused on the still strong π-withdrawing formyl (-CHO) and carboxyl (-COOH) groups, which also preserve the intra H-bond between the C5 substituent and the N4 amino group of Z, see Figure 1. The formyl and carboxyl moieties, in particular C5 atom modifications of cytosine, have also been very well studied in case of epigenetic modifications, and in context of DNA structures.[57,58] The resulting base pairs are named 5-CHO-Z:P and 5-COOH-Z:P in the subsequent text. Also in this case the optimized geometry of the 5-CHO-Z:P and 5-COOH-Z:P base pairs is very similar to that of the parent Z:P base pair, with only a slight elongation of the N4-H(Z)…O6(P) H-bond by 0.05 Å. However, the interaction strength of both the 5-CHO-Z:P and 5-COOH-Z:P base pairs are consistently lower, with E_{mod} of +1.26 and +1.52 kcal/mol in gas phase and +0.57 and +0.66 kcal/mol in water, respectively.

Next, to have a more comprehensive picture, we also investigated a series of halogens (fluoro, chloro and bromo) on the C5 atom of the Z base, again to examine the influence of functional groups that can differently impact the electron density on the Z-base, and, in turn, on the Z:P base pair. Finally, the analysis is completed by considering the impact of the strong electron donating –OMe and amino (-NH_2) groups, as these two strongly electron donating functional groups are at the opposite extremes of the Hammet scale, relative to the nitro group.[51]

Focusing on the optimized geometry of base pairs including halogen substituents, a systematic slight elongation of the N4(Z)-O6(P) H-bond, within 0.08 Å, is observed as compared to the Z:P pair.

Moving to gas phase energies, a significant destabilization effect, with E_{mod} within +1.99 kcal/mol, is observed for the base pairs presenting a halogen at the C5 position of the Z base. However, due to solvent screening effects, the E_{mod} is reduced to +0.34 kcal/mol in water calculations, again indicating a destabilization effect as compared to Z:P base pair. Similarly to the halogen substituents, in the 5-NH2-Z:P pair, the N4(Z)-O6(P) H-bond is elongated by 0.10 Å in gas phase, accompanied by a large destabilization effect, with E_{mod} of +3.12 kcal/mol and +0.78 kcal/mol in gas phase and in water, respectively.

Since the above analyses indicated that electron withdrawing groups at the C5 position stabilize the Z:P pair, whereas electron donating groups have a destabilizing effect, we decided to plot the Interaction energy versus the Hammet constant σ_P, in order to verify if an insightful relation between Interaction energy and σ_P could be obtained. Gratifyingly, we found a remarkable linear correlation, with $R^2 = 0.83$, over the whole Hammett scale, see Figure 5, clearly indicating the fundamental role of the nitro group in the original Z base to confer the Z:P pair a stability comparable to the natural G:C pair. Further, the plot of Figure 5 can be used to foresee the impact of different substituents if modifications of the original Z:P pair are designed.

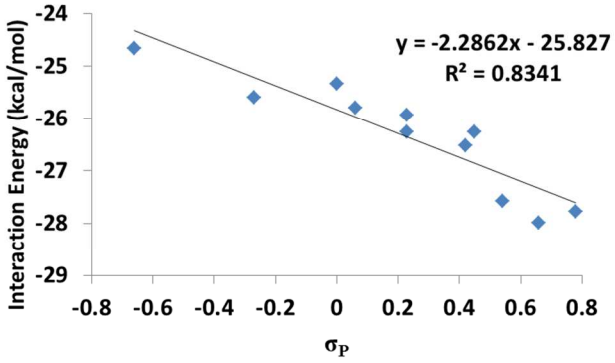


Figure 5. Plot showing the substituent effect, defined by Hammett constant (σ_P) versus the interaction energy of the base pairs under study.

Effect of Z-Nitro group on base-base stacking Interactions in A-DNA structure:

Visual inspection of the X-ray structure with PDB ID: 4XNO clearly indicated that the NO_2 group of Z in A-DNA facilitates stacking interactions with the purine or pyrimidine ring of the nucleobase adjacent to it, probably contributing to the stability of DNA in the A-form.[32] To support this proposal, we decided to quantify the role of the nitro group of the Z base on the stacking of Z with

other bases. To this end, we evaluated the relative base-base stacking energy (annotated as X//Y) of Z//P and Z//Z, and we compared it to the base-base stacking energy of 5-H-Z//P and 5-H-Z//Z. Since optimization of stacked bases often lead to too large geometrical distortions,[77-80] particularly in the absence of the nitro group, we preferred to freeze the geometry as in the crystallographic structure. For this reason, geometry optimization was confined to optimization of the hydrogen atoms only. The final geometries are shown in Figure 6. The main result we obtained is that the Z//P stacked geometry is clearly more stable both in gas phase and in water, by -2.12 and -1.61 kcal/mol, than the 5-H-Z//P geometry, which strongly supports the key role played by the Z-NO_2 group in inducing this stacking geometry. Similarly, the Z//Z geometry is more stable, by -1.61 and -2.48 kcal/mol in gas phase and in water, respectively, than the 5-H-Z//Z geometry. To conclude, our preliminary calculations match very well the experimental observation of preferential stacking of the Z-nitro group with adjacent purine or pyrimidine rings, which, in turn, would result in an increased stability of the A-DNA structure.

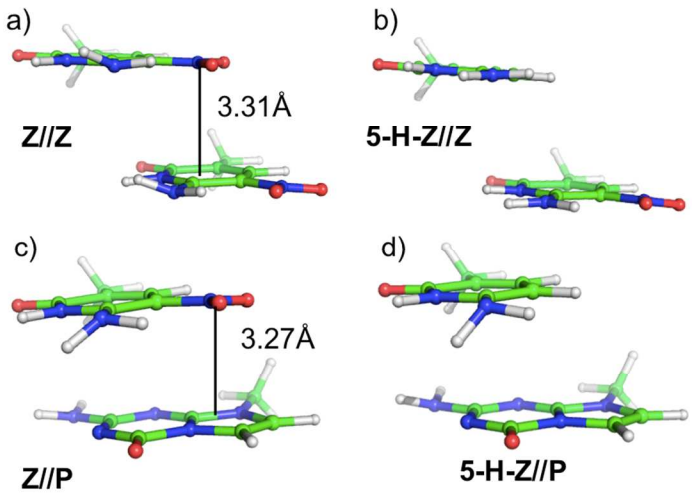


Figure 6. Structure of the stacked base//base pairs, where the Z-NO2 is stacked over the Z (a) and the P base (c), emphasized by the solid black line, and stacking interaction of the modeled 5-H-Z-base (nitro group replaced by a hydrogen atom) with the Z (b) and the P base (d).

Conclusions

In this work we have examined the stability of the non-natural Z:P base pair in the gas-phase and in water, and we found it to be comparable to the stability of the G:C pair. Further, the C1'-C1' distance of the freely optimized Z:P pair substantially matches the C1'-C1' distance of the freely optimized G:C pair. These results allow concluding that the Z:P pair closely mimics the G:C pair in terms of geometry and stability. To dissect the role of the nitro group on the geometry and stability

of the Z:P base pair, we compared a series of substituents at the C5 position of the Z base, having different electron donor/withdrawing properties. According to our calculations, the functional group on C5 of the Z base has a substantial impact on the Z:P base pair, with electron withdrawing groups stabilizing the H-bonded Z:P base pair. Indeed, changing the nature of the Z group at the C5 position from NO_2 to NH_2 allows tuning the stability of the Z:P base pair in a window of 4 kcal/mol. Finally, our calculations confirm that the Z-nitro group facilitates stacking interaction with the aromatic ring of adjacent purine or pyrimidine bases, thus imparting to the Z base unique structural properties DNA. In conclusion, our work allows cataloguing the synthetic Z:P base pair as a close mimic of the G:C base pair. Overall, these properties clarify the ability of the Z:P pair to effectively be used as another coding element in a six nucleotide genetic alphabet.

Acknowledgments.

The research reported in this publication was supported by funding from King Abdullah University of Science and Technology (KAUST).

Supporting Information.

Cartesian coordinates of all the structures discussed in this work. This material is available free of charge via the Internet at http://pubs.acs.org.

References:

(1) Benner, S. A.; Sismour, A. M. Synthetic biology. *Nat. Rev. Genet.* **2005**, *6*, 533-543.
(2) Benner, S. A. Redesigning life - organic-chemistry and the evolving protein. *Chimia* **1987**, *41*, 142-148.
(3) Szostak, J. W.; Bartel, D. P.; Luisi, P. L. Synthesizing life. *Nature* **2001**, *409*, 387-390.
(4) Hikida, Y.; Kimoto, M.; Yokoyama, S.; Hirao, I. Site-Specific fluorescent probing of rna molecules by unnatural base-pair transcription for local structural conformation analysis. *Nat. Protoc.* **2010**, *5*, 1312-1323.
(5) Kimoto, M.; Mitsui, T.; Yamashige, R.; Sato, A.; Yokoyama, S.; Hirao, I. A new unnatural base pair system between fluorophore and quencher base analogues for nucleic acid-based imaging technology. *J. Am. Chem. Soc.* **2010**, *132*, 15418-15426.
(6) Kimoto, M.; Mitsui, T.; Yokoyama, S.; Hirao, I. A unique fluorescent base analogue for the expansion of the genetic alphabet. *J. Am. Chem. Soc.* **2010**, *132*, 4988-4989.
(7) Hoshika, S.; Chen, F.; Leal, N. A.; Benner, S. A. Artificial genetic systems: self-avoiding dna in pcr and multiplexed pcr. *Angew. Chem. Int. Ed.* **2010**, *49*, 5554-5557.
(8) Sherrill, C. B.; Marshall, D. J.; Moser, M. J.; Larsen, C. A.; Daude-Snow, L.; Jurczyk, S.; Shapiro, G.; Prudent, J. R. Nucleic acid analysis using an expanded genetic alphabet to quench fluorescence. *J. Am. Chem. Soc.* **2004**, *126*, 4550-4556.

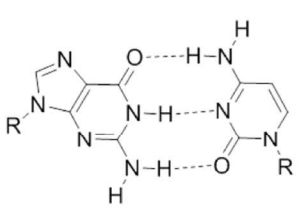
(9) Hirao, I.; Kimoto, M.; Mitsui, T.; Fujiwara, T.; Kawai, R.; Sato, A.; Harada, Y.; Yokoyama, S. An unnatural hydrophobic base pair system: site-specific incorporation of nucleotide analogs into dna and rna. *Nat. Methods* **2006**, *3*, 729-735.
(10) Kimoto, M.; Endo, M.; Mitsui, T.; Okuni, T.; Hirao, I.; Yokoyama, S. Site-specific incorporation of a photo-crosslinking component into rna by t7 transcription mediated by unnatural base pairs. *Chem. Biol.* **2004**, *11*, 47-55.
(11) Kimoto, M.; Kawai, R.; Mitsui, T.; Yokoyama, S.; Hirao, I. An unnatural base pair system for efficient pcr amplification and functionalization of dna molecules. *Nucleic Acids Res.* **2009**, *37*, e14.
(12) Kimoto, M.; Mitsui, T.; Harada, Y.; Sato, A.; Yokoyama, S.; Hirao, I. Fluorescent probing for rna molecules by an unnatural base-pair system. *Nucleic Acids Res.* **2007**, *35*, 5360-5369.
(13) Hirao, I. Unnatural base pair systems for dna/rna-based biotechnology. *Curr. Opin. Chem. Biol.* **2006**, *10*, 622-627.
(14) Hirao, I. Placing extra components into rna by specific transcription using unnatural base pair systems. *Biotechniques* **2006**, *40*, 711, 713, 715 passim.
(15) Hirao, I.; Ohtsuki, T.; Fujiwara, T.; Mitsui, T.; Yokogawa, T.; Okuni, T.; Nakayama, H.; Takio, K.; Yabuki, T.; Kigawa, T. et al. An unnatural base pair for incorporating amino acid analogs into proteins. *Nat. Biotechnol.* **2002**, *20*, 177-182.
(16) Henry, A. A.; Olsen, A. G.; Matsuda, S.; Yu, C.; Geierstanger, B. H.; Romesberg, F. E. Efforts to expand the genetic alphabet: identification of a replicable unnatural dns self-pair. *J. Am. Chem. Soc.* **2004**, *126*, 6923-6931.
(17) Henry, A. A.; Romesberg, F. E. Beyond a, c, g and t: augmenting nature's alphabet. *Curr. Opin. Chem. Biol.* **2003**, *7*, 727-733.
(18) Henry, A. A.; Yu, C.; Romesberg, F. E. Determinants of unnatural nucleobase stability and polymerase recognition. *J. Am. Chem. Soc.* **2003**, *125*, 9638-9646.
(19) Leconte, A. M.; Hwang, G. T.; Matsuda, S.; Capek, P.; Hari, Y.; Romesberg, F. E. Discovery, characterization, and optimization of an unnatural base pair for expansion of the genetic alphabet. *J. Am. Chem. Soc.* **2008**, *130*, 2336-2343.
(20) Leconte, A. M.; Matsuda, S.; Hwang, G. T.; Romesberg, F. E. Efforts towards expansion of the genetic alphabet: pyridone and methyl pyridone nucleobases. *Angew. Chem. Int. Ed.* **2006**, *45*, 4326-4329.
(21) Matsuda, S.; Fillo, J. D.; Henry, A. A.; Rai, P.; Wilkens, S. J.; Dwyer, T. J.; Geierstanger, B. H.; Wemmer, D. E.; Schultz, P. G.; Spraggon, G.; Romesberg, F. E. Efforts toward expansion of the genetic alphabet: structure and replication of unnatural base pairs. *J. Am. Chem. Soc.* **2007**, *129*, 10466-10473.
(22) Sismour, A. M.; Benner, S. A. The use of thymidine analogs to improve the replication of an extra dna base pair: a synthetic biological system. *Nucleic Acids Res.* **2005**, *33*, 5640-5646.
(23) Tae, E. L.; Wu, Y.; Xia, G.; Schultz, P. G.; Romesberg, F. E. Efforts toward expansion of the genetic alphabet: replication of dna with three base pairs. *J. Am. Chem. Soc.* **2001**, *123*, 7439-7440.
(24) Morales, J. C.; Kool, E. T. Minor groove interactions between polymerase and dna: more essential to replication than Watson-Crick hydrogen bonds? *J. Am. Chem. Soc.* **1999**, *121*, 2323-2324.
(25) Guckian, K. M.; Krugh, T. R.; Kool, E. T. Solution structure of a dna duplex containing a replicable difluorotoluene-adenine pair. *Nat. Struct. Biol.* **1998**, *5*, 954-959.
(26) Morales, J. C.; Kool, E. T. Efficient replication between non-hydrogen-bonded nucleoside shape analogs. *Nat. Struct. Biol.* **1998**, *5*, 950-954.
(27) Moran, S.; Ren, R. X.; Kool, E. T. A thymidine triphosphate shape analog lacking Watson-Crick pairing ability is replicated with high sequence selectivity. *Proc. Natl. Acad. Sci. U S A* **1997**, *94*, 10506-10511.
(28) Betz, K.; Malyshev, D. A.; Lavergne, T.; Welte, W.; Diederichs, K.; Dwyer, T. J.; Ordoukhanian, P.; Romesberg, F. E.; Marx, A. KlenTaq polymerase replicates unnatural base pairs by inducing a Watson-Crick geometry. *Nat. Chem. Biol.* **2012**, *8*, 612-614.
(29) Malyshev, D. A.; Dhami, K.; Lavergne, T.; Chen, T.; Dai, N.; Foster, J. M.; Correa, I. R., Jr.; Romesberg, F. E. A semi-synthetic organism with an expanded genetic alphabet. *Nature* **2014**, *509*, 385-388.
(30) Malyshev, D. A.; Seo, Y. J.; Ordoukhanian, P.; Romesberg, F. E. PCR with an expanded genetic alphabet. *J. Am. Chem. Soc.* **2009**, *131*, 14620-14621.

(31) Zhang, L.; Yang, Z.; Sefah, K.; Bradley, K. M.; Hoshika, S.; Kim, M. J.; Kim, H. J.; Zhu, G.; Jimenez, E.; Cansiz, et al. Evolution of functional six-nucleotide dna. *J. Am. Chem. Soc.* **2015**, *137*, 6734-6737.
(32) Georgiadis, M. M.; Singh, I.; Kellett, W. F.; Hoshika, S.; Benner, S. A.; Richards, N. G. Structural basis for a six nucleotide genetic alphabet. *J. Am. Chem. Soc.* **2015**, *137*, 6947-6955.
(33) Hernandez, A. R.; Shao, Y.; Hoshika, S.; Yang, Z.; Shelke, S. A.; Herrou, J.; Kim, H. J.; Kim, M. J.; Piccirilli, J. A.; Benner, S. A. A crystal structure of a functional rna molecule containing an artificial nucleobase pair. *Angew. Chem., Int. Ed.* **2015**, *54*, 9853-9856.
(34) Benner, S. A. Understanding nucleic acids using synthetic chemistry. *Acc. Chem. Res.* **2004**, *37*, 784-797.
(35) Chawla, M.; Abdel-Azeim, S.; Oliva, R.; Cavallo, L. Higher order structural effects stabilizing the reverse watson-crick guanine-cytosine base pair in functional rnas. *Nucleic Acids Res.* **2014**, *42*, 714-726.
(36) Chawla, M.; Credendino, R.; Oliva, R.; Cavallo, L. Structural and energetic impact of non-natural 7-deaza-8-azaadenine and its 7-substituted derivatives on h-Bonding potential with uracil in rna molecules. *J. Phys. Chem. B* **2015**, *119*, 12982-12989.
(37) Chawla, M.; Credendino, R.; Poater, A.; Oliva, R.; Cavallo, L. Structural stability, acidity, and halide selectivity of the fluoride riboswitch recognition site. *J. Am. Chem. Soc.* **2015**, *137*, 299-306.
(38) Chawla, M.; Oliva, R.; Bujnicki, J. M.; Cavallo, L. An atlas of rna base pairs involving modified nucleobases with optimal geometries and accurate energies. *Nucleic Acids Res.* **2015**, *43*, 6714-6729.
(39) Chawla, M.; Sharma, P.; Halder, S.; Bhattacharyya, D.; Mitra, A. Protonation of base pairs in rna: context analysis and quantum chemical investigations of their geometries and stabilities. *J. Phys. Chem. B* **2011**, *115*, 1469-1484.
(40) Oliva, R.; Cavallo, L. Frequency and effect of the binding of Mg^{2+}, Mn^{2+}, and Co^{2+} ions on the guanine base in watson-crick and reverse watson-crick base pairs. *J. Phys. Chem. B* **2009**, *113*, 15670-15678.
(41) Oliva, R.; Cavallo, L.; Tramontano, A. Accurate energies of hydrogen bonded nucleic acid base pairs and triplets in tRNA tertiary interactions. *Nucleic Acids Res.* **2006**, *34*, 865-879.
(42) Oliva, R.; Tramontano, A.; Cavallo, L. Mg^{2+} binding and archaeosine modification stabilize the g15 c48 levitt base pair in tRNAs. *RNA* **2007**, *13*, 1427-1436.
(43) Sharma, P.; Chawla, M.; Sharma, S.; Mitra, A. On the role of hoogsteen:hoogsteen interactions in rna: ab initio investigations of structures and energies. *RNA* **2010**, *16*, 942-957.
(44) Sharma, P.; Sharma, S.; Chawla, M.; Mitra, A. Modeling the noncovalent interactions at the metabolite binding site in purine riboswitches. *J. Mol. Model.* **2009**, *15*, 633-649.
(45) Sponer, J.; Jurecka, P.; Hobza, P. Accurate interaction energies of hydrogen-bonded nucleic acid base pairs. *J. Am. Chem. Soc.* **2004**, *126*, 10142-10151.
(46) Sponer, J.; Mladek, A.; Sponer, J. E.; Svozil, D.; Zgarbova, M.; Banas, P.; Jurecka, P.; Otyepka, M. The dna and rna sugar-phosphate backbone emerges as the key player. an overview of quantum-chemical, structural biology and simulation studies. *Phys. Chem. Chem. Phys.* **2012**, *14*, 15257-15277.
(47) Sponer, J. E.; Leszczynski, J.; Sychrovsky, V.; Sponer, J. Sugar edge/sugar edge base pairs in rna: stabilities and structures from quantum chemical calculations. *J. Phys. Chem. B* **2005**, *109*, 18680-18689.
(48) Sponer, J. E.; Reblova, K.; Mokdad, A.; Sychrovsky, V.; Leszczynski, J.; Sponer, J. Leading RNA tertiary interactions: structures, energies, and water insertion of a-minor and p-interactions. A quantum chemical view. *J. Phys. Chem. B* **2007**, *111*, 9153-9164.
(49) Sponer, J. E.; Spackova, N.; Kulhanek, P.; Leszczynski, J.; Sponer, J. Non-watson-crick base pairing in rna. quantum chemical analysis of the cis watson-crick/sugar edge base pair family. *J. Phys. Chem. A* **2005**, *109*, 2292-2301.
(50) Sponer, J. E.; Spackova, N.; Leszczynski, J.; Sponer, J. Principles of rna base pairing: structures and energies of the trans watson-Crick/sugar edge base pairs. *J. Phys. Chem. B* **2005**, *109*, 11399-11410.
(51) Hansch, C.; Leo, A.; Taft, R. W. A survey of hammett substituent constants and resonance and field parameters. *Chem. Rev.* **1991**, *91*, 165-195.

(52) Boyd, G. W.; Coombs, M. M.; Baird, W. M. The presence of a trifluoromethyl rather than a methyl substituent in the bay-region greatly decreases the dna-binding and tumour-initiating activity of the cyclopenta[alpha]phenanthren-17-ones. *Carcinogenesis* **1995**, *16*, 2543-2547.
(53) Satake, H.; Takeda, S.; Matsumura, A.; Sasaki, M.; Sugimoto, N.; Wataya, Y. Action of 5-trifluoromethyl-2'-deoxyuridine on dna synthesis. *Nucleic Acids Symp. Ser.* **1992**, 189-190.
(54) Taranenko, M.; Rykhlevskaya, A.; Mtchedlidze, M.; Laval, J.; Kuznetsova, S. Photochemical cross-linking of Escherichia coli fpg protein to dna duplexes containing phenyl(trifluoromethyl)diazirine groups. *Eur. J. Biochem.* **2003**, *270*, 2945-2949.
(55) Watson, M. D.; Gai, X. S.; Gillies, A. T.; Brewer, S. H.; Fenlon, E. E. A vibrational probe for local nucleic acid environments: 5-cyano-2'-deoxyuridine. *J. Phys. Chem. B* **2008**, *112*, 13188-13192.
(56) Kim, H. J.; Chen, F.; Benner, S. A. Synthesis and properties of 5-cyano-substituted nucleoside analog with a donor-donor-acceptor hydrogen-bonding pattern. *J Org. Chem.* **2012**, *77*, 3664-3669.
(57) Bachman, M.; Uribe-Lewis, S.; Yang, X.; Burgess, H. E.; Iurlaro, M.; Reik, W.; Murrell, A.; Balasubramanian, S. 5-Formylcytosine can be a stable dna modification in mammals. *Nat. Chem. Biol.* **2015**, *11*, 555-557.
(58) Kellinger, M. W.; Song, C. X.; Chong, J.; Lu, X. Y.; He, C.; Wang, D. 5-formylcytosine and 5-carboxylcytosine reduce the rate and substrate specificity of rna polymerase II transcription. *Nat. Struct. Mol. Biol.* **2012**, *19*, 831-833.
(59) Phelps, K. J.; Ibarra-Soza, J. M.; Tran, K.; Fisher, A. J.; Beal, P. A. Click modification of rna at adenosine: structure and reactivity of 7-ethynyl- and 7-triazolyl-8-aza-7-deazaadenosine in rna. *ACS Chem. Biol.* **2014**, *9*, 1780-1787.
(60) Seela, F.; He, Y.; Wei, C. F. Parallel-stranded oligonucleotide duplexes containing 5-methylisocytosine-guanine and isoguanine-cytosine base pairs. *Tetrahedron* **1999**, *55*, 9481-9500.
(61) Seela, F.; Wei, C. F.; Melenewski, A.; Feiling, E. Parallel-stranded duplex dna and self-assembled quartet structures formed by isoguanine and related bases. *Nucleosides, Nucleotides Nucleic Acids* **1998**, *17*, 2045-2052.
(62) Seela, F.; Wei, C. F.; Melenewski, A.; He, Y.; Kroschel, R.; Feiling, E. Parallel-stranded dna formed by new base pairs related to the isoguanine-cytosine or isocytosine-guanine motifs. *Nucleosides, Nucleotides Nucleic Acids* **1999**, *18*, 1543-1548.
(63) Sugiyama, H.; Ikeda, S.; Saito, I. Remarkably stable parallel-stranded oligonucleotides containing 5-methylisocytosine and isoguanine. *J. Am. Chem. Soc,* **1996**, *118*, 9994-9995.
(64) Yang, X. L.; Sugiyama, H.; Ikeda, S.; Saito, I.; Wang, A. H. Structural studies of a stable parallel-stranded DNA duplex incorporating isoguanine:cytosine and isocytosine:guanine basepairs by nuclear magnetic resonance spectroscopy. *Biophys. J.* **1998**, *75*, 1163-1171.
(65) Becke, A. D. Density-Functional Thermochemistry .3. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648-5652.
(66) Becke, A. D. Density-functional thermochemistry. *Abst. Pap. Am. Chem. Soc.* **1996**, *212*.
(67) Lee, C. T.; Yang, W. T.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron-density. *Phys. Rev. B* **1988**, *37*, 785-789.
(68) Dunning, T. H. Gaussian-Basis sets for use in correlated molecular calculations .1. the atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007-1023.
(69) Klamt, A.; Schuurmann, G. Cosmo - a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc.,Perkin Trans. 2* **1993**, 799-805.
(70) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (dft-d) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*.
(71) Moller, C.; Plesset, M. S. Note on an approximation treatment for many-electron systems. *Phys. Rev.* **1934**, *46*, 0618-0622.
(72) Weigend, F.; Haser, M. RI-MP2: first derivatives and global consistency. *Theor. Chem. Acc.* **1997**, *97*, 331-340.
(73) Boys, S. F.; Bernardi, F. Calculation of small molecular interactions by differences of separate total energies - some procedures with reduced errors. *Mol. Phys.* **1970**, *19*, 553.

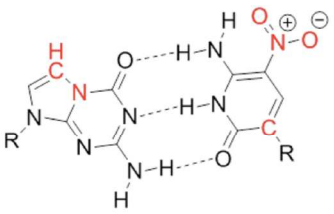
(74) Zirbel, C. L.; Sponer, J. E.; Sponer, J.; Stombaugh, J.; Leontis, N. B. Classification and energetics of the base-phosphate interactions in rna. *Nucleic Acids Res.* **2009**, *37*, 4898-4918.
(75) Leontis, N. B.; Stombaugh, J.; Westhof, E. The non-watson-crick base pairs and their associated isostericity matrices. *Nucleic Acids Res.* **2002**, *30*, 3497-3531.
(76) Leontis, N. B.; Westhof, E. Geometric nomenclature and classification of rna base pairs. *RNA* **2001**, *7*, 499-512.
(77) Hobza, P.; Sponer, J. Significant structural deformation of nucleic acid bases in stacked base pairs: an ab initio study beyond Hartree-Fock. *Chem. Phys. Lett.* **1998**, *288*, 7-14.
(78) McDonald, A. R.; Denning, E. J.; MacKerell, A. D. Impact of geometry optimization on base-base stacking interaction energies in the canonical a- and b-forms of dna. *J. Phys. Chem. A* **2013**, *117*, 1560-1568.
(79) Sponer, J.; Leszczynski, J.; Hobza, P. Nature of nucleic acid-base stacking: Nonempirical ab initio and empirical potential characterization of 10 stacked base dimers. comparison of stacked and h-bonded base pairs. *J. Phys. Chem.* **1996**, *100*, 5590-5596.
(80) Sponer, J.; Leszczynski, J.; Hobza, P. Hydrogen bonding and stacking of dna bases: a review of quantum-chemical ab initio studies. *J. Biomol. Struct. Dyn.* **1996**, *14*, 117-135.

Material for the Table of Contents



G:C cWW

Versus...



P:Z cWW